

# Overcrowding asymptotics for the Sine $_{\beta}$ process

Diane Holcomb\*      Benedek Valkó†

June 24, 2015

## Abstract

We give overcrowding estimates for the Sine $_{\beta}$  process, the bulk point process limit of the Gaussian  $\beta$ -ensemble. We show that the probability of having at least  $n$  points in a fixed interval is given by  $e^{-\frac{\beta}{2}n^2 \log(n) + \mathcal{O}(n^2)}$  as  $n \rightarrow \infty$ . We also identify the next order term in the exponent if the size of the interval goes to zero.

## 1 Introduction

The Gaussian  $\beta$ -ensemble is a family of point processes on the real line given by the joint density function

$$p_{m,\beta}(\lambda_1, \dots, \lambda_m) = \frac{1}{Z_{m,\beta}} \prod_{1 \leq i < j \leq m} |\lambda_i - \lambda_j|^{\beta} \times \prod_{i=1}^m e^{-\frac{\beta}{4}\lambda_i^2}. \quad (1)$$

Here  $\beta > 0$  and  $Z_{m,\beta}$  is an explicitly computable normalizing constant. For the specific values of  $\beta = 1, 2$  and  $4$  one obtains the joint eigenvalue density of the classical random matrix ensembles GOE, GUE and GSE.

We study the Sine $_{\beta}$  process, the bulk scaling limit of the Gaussian  $\beta$ -ensemble. Fix  $\beta > 0$ , and consider a sequence of finite point processes  $\Lambda_{m,\beta}$  with distribution given by the density (1). Then as  $m \rightarrow \infty$  the scaled processes  $2\sqrt{m} \Lambda_{m,\beta}$  converge in distribution to a translation invariant point process which we call Sine $_{\beta}$ . The existence and description of the limiting process in the  $\beta = 1, 2$  and  $4$  cases were obtained by Dyson, Gaudin and Mehta (see e.g. the monographs [1], [9], [13]), while in the general  $\beta$  case this was done in [14]. We note that in our normalization the Sine $_{\beta}$  process has particle density  $\frac{1}{2\pi}$ .

---

\*Department of Mathematics, University of Arizona, dianeholcomb@math.arizona.edu

†Department of Mathematics, University of Wisconsin - Madison, valko@math.wisc.edu

We study the problem of overcrowding for the Sine $_{\beta}$  process. We describe the asymptotic probability of having at least  $n$  points in a fixed interval for a large  $n$ . Since the process is translation invariant we may assume that the interval is of the form  $[0, \lambda]$  with  $\lambda > 0$ . We will use the notation  $N_{\beta}(\lambda)$  for the counting function of the process, for  $\lambda > 0$  this gives the number of points in  $[0, \lambda]$ . Our main result is as follows.

**Theorem 1.** *Fix  $\lambda_0 > 0$ . Then there is a constant  $c > 0$  depending only on  $\beta$  and  $\lambda_0$  so that for any  $n \geq 1$  and  $0 < \lambda \leq \lambda_0$  we have*

$$e^{-\frac{\beta}{2}n^2 \log\left(\frac{n}{\lambda}\right) - cn \log(n+1) \log\left(\frac{n}{\lambda}\right) - cn^2} \leq P(N_{\beta}(\lambda) \geq n) \leq e^{-\frac{\beta}{2}n^2 \log\left(\frac{n}{\lambda}\right) + cn \log(n+1) \log\left(\frac{n}{\lambda}\right) + cn^2}. \quad (2)$$

The theorem implies that for a fixed  $\lambda > 0$  we have  $P(N_{\beta}(\lambda) \geq n) = e^{-\frac{\beta}{2}n^2 \log n + \mathcal{O}(n^2)}$  as  $n \rightarrow \infty$ . If we let  $n \rightarrow \infty$  and  $\lambda \rightarrow 0$  simultaneously then we get

$$P(N_{\beta}(\lambda) \geq n) = e^{-\frac{\beta}{2}n^2 \log(n/\lambda) + \mathcal{O}(n^2 + n \log n \log(\lambda^{-1}))}.$$

Dyson used Coulomb gas arguments in [6] to predict the asymptotic size of  $P(N_{\beta}(\lambda) = n)$  for a fixed  $n$  and large  $\lambda$ . His predictions were made precise in [5] for  $\beta = 1, 2$  and  $4$  using the fact that in these classical cases the probabilities in question have Fredholm determinant representations. For general  $\beta$  the special case of  $n = 0$  (i.e. the large gap probability) was treated in [15].

Fogler and Shklovskii in [8] extended Dyson's method to give predictions on the asymptotic size of  $P(N_{\beta}(\lambda) = n)$  in the regime where  $n$  and  $\lambda$  are large. Although this does not cover the case we are interested in ( $\lambda$  bounded,  $n$  large), our results agree with their prediction in the case when  $n \gg \lambda$ . Section 14.6 of [9] gives a nice overview of the Coloumb gas method in the study of the distribution of  $N_{\beta}(\lambda)$ .

Our proof uses techniques that we developed in [10], where we proved a large deviation principle for  $\frac{1}{\lambda}N_{\beta}(\lambda)$  with scale  $\lambda^2$  and a good rate function of the form  $\beta I(\rho)$ . One of the consequences of this large deviation result is that if  $\rho > \frac{1}{2\pi}$  then  $P(N_{\beta}(\lambda) \geq \rho\lambda)$  decays asymptotically as  $e^{-\beta\lambda^2 I(\rho)(1+o(1))}$  as  $\lambda \rightarrow \infty$ . It was shown in [10] that  $I(\rho)$  grows like  $\frac{1}{2}\rho^2 \log \rho$  as  $\rho \rightarrow \infty$  so the decay of  $P(N_{\beta}(\lambda) \geq \rho\lambda)$  is formally consistent with our Theorem 1, even though it describes a different regime of  $n, \lambda$ .

Rigorous overcrowding estimates have been obtained for certain non-Hermitian random matrix models. Here one is interested in the asymptotic probability of finding at least  $n$  points in a fixed disk. In [11] Krishnapur identified the overcrowding asymptotics for the complex Ginibre ensemble and also gave similar estimates for zeros of certain Gaussian analytic functions. In [3] and [2] these results were generalized to point processes obtained

by considering products of complex and quaternion Ginibre ensembles. It is interesting to note that in all of these non-Hermitian cases the probability of finding at least  $n$  points in a fixed disk decays like  $e^{-cn^2 \log n(1+o(1))}$  where  $c$  is a parameter depending only on the model and not the size of the disk. This is of the same form as our result for the  $\text{Sine}_\beta$  process.

The finite version of our problem would be to study the probability of having a large number of points in a small interval of order  $m^{-1/2}$  in the bulk of the support of (1). This problem has been considered for the spectrum of general Wigner matrices in the context of showing the universality of the bulk limit process, see e.g. Theorem 3.5 of [7] or Corollary B.2 of [4]. These results concentrate on giving sufficiently strong (but not necessarily optimal) upper bounds for the overcrowding probabilities.

The starting point of our proof is the following characterization of the counting function of  $\text{Sine}_\beta$ , due to [14].

**Theorem 2** ([14]). *Let  $\lambda > 0$  and consider the strong solution of the following diffusion:*

$$d\alpha_\lambda = \lambda \frac{\beta}{4} e^{-\frac{\beta}{4}t} dt + 2 \sin(\alpha_\lambda/2) dB_t, \quad \alpha_\lambda(0) = 0. \quad (3)$$

*Then  $\lim_{t \rightarrow \infty} \frac{1}{2\pi} \alpha_\lambda(t)$  exists almost surely, and it has the same distribution as  $N_\beta(\lambda)$ .*

This describes the one-dimensional marginals of the counting function  $N_\beta(\lambda)$ . One can also describe the distribution of the whole function  $N_\beta(\cdot)$  in terms of a one-parameter family of diffusions (see [14]), but we will not need this here. The following proposition summarizes some of the basic properties of the diffusion (3). Most of these properties are straightforward, see Proposition 9 in [14] for some of the proofs.

**Proposition 3.** *Let  $\lambda > 0$ .*

- (i) *The process  $\lfloor \frac{\alpha_\lambda(t)}{2\pi} \rfloor$  is nondecreasing in  $t$ . In particular for any integer  $n$  we have  $\alpha_\lambda(t) > 2\pi n$  if  $t$  is larger than the first hitting time of  $2\pi n$ .*
- (ii) *If  $0 < \lambda < \lambda'$  and  $\alpha_\lambda, \alpha_{\lambda'}$  are strong solutions of (3) with the same Brownian motion  $B_t$  then  $\alpha_\lambda(t) \leq \alpha_{\lambda'}(t)$  for all  $t \geq 0$ .*
- (iii) *For any integer  $n$  the process  $\alpha_\lambda(t) + 2n\pi$  solves the same SDE as (3) but with an initial condition of  $2n\pi$  instead of 0.*
- (iv) *For any  $T > 0$  the process  $\phi(t) = \alpha_\lambda(T+t)$  solves the same SDE as (3) with  $\lambda e^{-\frac{\beta}{4}T}$  in place of  $\lambda$ ,  $B_{t+T}$  in place of  $B_t$  and initial condition  $\phi(0) = \alpha_\lambda(T)$ .*

An important consequence of the previous proposition is the following observation.

**Lemma 4.** *Fix  $n > 1$  and  $\lambda > 0$ . Let  $\tau_\lambda$  be the first hitting time of  $2\pi$  of the diffusion  $\alpha_\lambda$ . (Note that  $\tau_\lambda$  might be  $\infty$ ). Then*

$$P(N_\beta(\lambda) \geq n) = E[g(\tau_\lambda)], \quad \text{where } g(t) = P\left(N_\beta(\lambda e^{-\frac{\beta}{4}t}) \geq n - 1\right), \quad g(\infty) = 0. \quad (4)$$

*Proof.* According to Proposition 3, if we condition on  $\{\tau_\lambda = T\}$  with a  $T > 0$  then the process  $\{\alpha_\lambda(t + T) - 2\pi, t \geq 0\}$  will have the same distribution as  $\{\alpha_{\lambda e^{-\frac{\beta}{4}T}}(t), t \geq 0\}$ . Thus

$$P\left(\lim_{t \rightarrow \infty} \alpha_\lambda(t) \geq 2\pi n \mid \tau_\lambda = T\right) = P\left(\lim_{t \rightarrow \infty} \alpha_{\lambda e^{-\frac{\beta}{4}T}}(t) \geq 2\pi(n - 1)\right). \quad (5)$$

The statement of the lemma now follows by taking expectations in  $\tau_\lambda$  and using the description of  $N_\beta(\lambda)$  from Theorem 2.  $\square$

We record the following simple corollary of Lemma 4.

**Proposition 5.** *For any  $\lambda > 0$  and  $n \geq 1$  we have*

$$P(N_\beta(\lambda) \geq n) \leq \left(\frac{\lambda}{2\pi}\right)^n. \quad (6)$$

*Proof.* From (3) one gets  $E\alpha_\lambda(t) = \lambda(1 - e^{-\frac{\beta}{4}t})$  and using the properties of the diffusion  $\alpha_\lambda$  one can show  $E \lim_{t \rightarrow \infty} \alpha_\lambda(t) = \lambda$ . Markov's inequality now implies the statement of the proposition for  $n = 1$ . For larger values we can use Lemma 4 and the monotonicity of  $g(t)$  from (4) to get

$$P(N_\beta(\lambda) \geq n) \leq P(\tau_\lambda < \infty)P(N_\beta(\lambda) \geq n - 1) \leq P(N_\beta(\lambda) \geq 1)P(N_\beta(\lambda) \geq n - 1)$$

from which the statement follows by induction.  $\square$

Our proof of Theorem 1 will rely on the recursion equation (4) and estimates on  $P(\tau_\lambda < t)$ . We will show that if we can get a good enough upper or lower bound on  $P(N_\beta(\lambda) \geq n - 1)$  for all  $\lambda \leq \lambda_0$ , then we can also get good bounds for  $P(N_\beta(\lambda) \geq n)$  for all  $\lambda \leq \lambda_0$ . The next section will contain the necessary hitting time estimates while the remaining sections will contain the proofs of the upper and lower bounds in Theorem 1.

## 2 Hitting time estimates

We will estimate the hitting time  $\tau_\lambda$  by coupling the diffusion  $\alpha_\lambda(t)$  of (3) to a similar diffusion with a constant drift term. Let  $\tilde{\alpha}_\lambda$  be the strong solution of the SDE

$$d\tilde{\alpha}_\lambda = \lambda dt + 2 \sin\left(\frac{\tilde{\alpha}_\lambda}{2}\right) dB_t, \quad \tilde{\alpha}_\lambda(0) = 0, \quad (7)$$

using the same Brownian motion as (3). Denote by  $\tilde{\tau}_\lambda$  the first hitting time of  $2\pi$  by  $\tilde{\alpha}_\lambda$ . Simple coupling arguments show that  $\alpha_\lambda(t) \leq \tilde{\alpha}_{\lambda'}(t)$  for all  $t \geq 0$  with  $\lambda' = \frac{\beta}{4}\lambda$ , and  $\alpha_\lambda(t) \geq \tilde{\alpha}_{\lambda''}(t)$  for  $0 \leq t \leq T$  if  $\lambda'' = \frac{\beta}{4}e^{-\frac{\beta}{4}T}\lambda$ . This yields

$$P(\tau_\lambda \leq t) \leq P(\tilde{\tau}_{\lambda'} \leq t), \quad \text{with} \quad \lambda' = \frac{\beta}{4}\lambda, \quad \text{and} \quad (8)$$

$$P(\tau_\lambda \leq t) \geq P(\tilde{\tau}_{\lambda''} \leq t), \quad \text{with} \quad \lambda'' = \frac{\beta}{4}e^{-\frac{\beta}{4}t}\lambda. \quad (9)$$

Our bounds will use certain special functions which we now briefly go over. For  $a < 1$  we set

$$K(a) = \int_0^{\pi/2} \frac{dx}{\sqrt{1 - a \sin^2 x}}, \quad E(a) = \int_0^{\pi/2} \sqrt{1 - a \sin^2 x} dx. \quad (10)$$

These functions are the complete elliptic integrals of the first and second kind. We note the identity

$$K(-a) = \frac{1}{2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\cosh^2 z + a}} dz, \quad a > -1. \quad (11)$$

We denote by  $\mathcal{W}$  the lower branch of the Lambert W function (or product log function). This is defined as the unique solution of

$$z = \mathcal{W}(z)e^{\mathcal{W}(z)} \quad \text{with} \quad \mathcal{W}(z) \leq -1. \quad (12)$$

Note that  $\mathcal{W}$  has domain  $[-1/e, 0]$ , it is strictly decreasing and satisfies  $\mathcal{W}(x \log x) = \log x$  for  $0 < x \leq 1/e$ .

The following proposition summarizes some of the relevant properties of the introduced special functions.

**Proposition 6.** *There exists a constant  $c > 0$ , so that the following bounds hold.*

$$\left| K(-a) - \frac{1}{2\sqrt{a}} \log(16a) \right| \leq \frac{c}{a^{3/2}} \log(a), \quad |E(-a) - \sqrt{a}| \leq \frac{c}{a^{1/2}} \log(a), \quad \text{if } a > 2. \quad (13)$$

$$|K^{-1}(x) + x^{-2}\mathcal{W}^2(-x/4)| \leq c, \quad \text{if } 0 < x < 1/2. \quad (14)$$

$$|\mathcal{W}(-x) - \mathcal{W}(-y)| \leq c |\log(x/y)|, \quad |\mathcal{W}(-x)| \leq c \log\left(\frac{1}{x}\right), \quad \text{if } x, y \in (0, \frac{1}{2e}]. \quad (15)$$

*Outline of the proof.* The estimates (13) are from the Appendix of [10]. The definition of  $\mathcal{W}$  allows us to check that the inverse of the function  $\frac{1}{2\sqrt{a}} \log(16a)$  on  $[e^2/16, \infty)$  is exactly  $x^{-2}\mathcal{W}^2(-x/4)$ . From this fact and the first bound in (13) the proof of (14) is a simple exercise. Finally, (15) follows from the observation that  $\mathcal{W}'(x) = \frac{1}{x} \cdot \frac{\mathcal{W}(x)}{1+\mathcal{W}(x)}$ .  $\square$

Now we are ready to state the estimates on the hitting times  $\tau_\lambda$  and  $\tilde{\tau}_\lambda$ .

**Proposition 7.** *There exist positive constants  $\varepsilon$  and  $c$  (depending only on  $\beta$ ) so that if  $\lambda t < \varepsilon$  then*

$$P(\tau_\lambda < t) \leq \exp \left[ -\frac{2}{t} \mathcal{W}^2(-\lambda t) + c \left( 1 + \frac{1}{t} \right) \log(\frac{1}{\lambda t}) \right]. \quad (16)$$

For any  $\lambda_0 > 0$  there exist positive constants  $\varepsilon, c$  (depending only on  $\lambda_0, \beta$ ) so that if  $0 < \lambda \leq \lambda_0$ ,  $\lambda t < \varepsilon$  and  $t \leq \varepsilon \log(\frac{1}{\lambda t})$  then

$$P(\tau_\lambda < t) \geq \exp \left[ -\frac{2}{t} \mathcal{W}^2(-\lambda t) - c \left( 1 + \frac{1}{t} \right) \log(\frac{1}{\lambda t}) \right]. \quad (17)$$

Moreover, the same statements hold for the hitting time  $\tilde{\tau}_\lambda$  (with possibly different constants  $\varepsilon$  and  $c$ ).

We will prove the upper and lower bounds separately. In both cases we will prove the estimate for  $\tilde{\tau}_\lambda$  first, then use (8) and (9) to get the corresponding statement for  $\tau_\lambda$ .

## 2.1 Proof of the lower bounds in Proposition 7

By introducing the new variable  $X_\lambda = \log(\tan(\tilde{\alpha}_\lambda/4))$  we can transform (7) into the SDE

$$dX_\lambda = \frac{\lambda}{2} \cosh X_\lambda dt + \frac{1}{2} \tanh X_\lambda dt + dB_t, \quad X_\lambda(0) = -\infty. \quad (18)$$

Under this transformation  $\tilde{\tau}_\lambda$  becomes the hitting time of  $\infty$  for the process  $X_\lambda$ . Following [10] we compare  $X_\lambda$  with another family of diffusions. For  $a \geq -1$  consider the SDE

$$dY_{\lambda,a} = \frac{\lambda}{2} \sqrt{\cosh^2 Y_{\lambda,a} + a} dt + \frac{1}{2} \tanh Y_{\lambda,a} dt + dB_t, \quad Y_{\lambda,a}(0) = -\infty, \quad (19)$$

and denote by  $\tau_{Y,\lambda,a}$  the hitting time of  $\infty$ . In [10] the Cameron-Martin-Girsanov formula was used for diffusions with explosions to compare  $X_\lambda$  and  $Y_{\lambda,a}$  and the following results were obtained.

**Proposition 8 ([10]).** *For  $a > -1$  and  $s_1 < s_2$  we have*

$$P(\tilde{\tau}_\lambda \in [s_1, s_2]) = E \left[ \mathbf{1}(\tau_{Y,\lambda,a} \in [s_1, s_2]) e^{-G_{\tau_{Y,\lambda,a}}(Y)} \right]. \quad (20)$$

Here  $G_{\tau_{Y,\lambda,a}}(Y)$  is the Girsanov factor for which the following estimate holds:

$$\left| G_{\tau_{Y,\lambda,a}}(Y) + \frac{\lambda^2 a \tau_{Y,\lambda,a}}{8} - \lambda((1+a)K(-a) - E(-a)) \right| \leq \frac{\lambda \sqrt{|a|}}{4} \tau_{Y,\lambda,a}. \quad (21)$$

The proof of these statements are contained in Section 3 of [10]. Note the slight notational inconsistency, in the statements in [10]  $-a$  is used in place of  $a$ .

Our lower bound proof will rely on the following lemma, which will be proved at the end of this subsection.

**Lemma 9.** *There is a constant  $c > 0$  so that for any  $a, \lambda$  with  $\lambda\sqrt{a} \geq 2$  and  $a > 2$  we have*

$$P\left(\lambda\tau_{Y,\lambda,a} \in \left[4K(-a)\left(1 - \frac{5}{\lambda\sqrt{a}}\right), 4K(-a)\left(1 + \frac{5}{\lambda\sqrt{a}}\right)\right]\right) \geq c\sqrt{K(-a)}. \quad (22)$$

Now we are ready to prove the lower bounds in Proposition 7.

*Proof of the lower bounds in Proposition 7.* We use (20) and (21) with an  $a > 0$  to get

$$\begin{aligned} P(\lambda\tilde{\tau}_\lambda \leq s_2) &\geq P(\lambda\tilde{\tau}_\lambda \in [s_1, s_2]) = E\left[\mathbf{1}(\lambda\tau_{Y,\lambda,a} \in [s_1, s_2])e^{-G_{\tau_{Y,\lambda,a}}(Y)}\right] \\ &\geq E\left[\mathbf{1}(\lambda\tau_{Y,\lambda,a} \in [s_1, s_2])e^{\frac{\lambda^2 a \tau_{Y,\lambda,a}}{8} - \lambda((1+a)K(-a) - E(-a)) - \frac{\lambda\sqrt{a}}{4}\tau_{Y,\lambda,a}}\right]. \end{aligned}$$

Assuming  $\lambda\sqrt{a} \geq 2$  we see that the exponential term under the last expectation is nondecreasing in  $\tau_{Y,\lambda,a}$  and because of the indicator function we get the bound

$$P(\lambda\tilde{\tau}_\lambda \leq s_2) \geq P(\lambda\tau_{Y,\lambda,a} \in [s_1, s_2])e^{\frac{\lambda a s_1}{8} - \lambda((1+a)K(-a) - E(-a)) - \frac{\sqrt{a}}{4}s_1}.$$

Now choose  $[s_1, s_2] = \left[4K(-a)\left(1 - \frac{5}{\lambda\sqrt{a}}\right), 4K(-a)\left(1 + \frac{5}{\lambda\sqrt{a}}\right)\right]$ , and assume  $a > 2$ . With Lemma 9 we can estimate the probability on the right and we can use the asymptotics (13) of Proposition 6 to estimate the exponential term. Collecting all the terms gives

$$P\left(\lambda\tilde{\tau}_\lambda \leq 4K(-a)\left(1 + \frac{5}{\lambda\sqrt{a}}\right)\right) \geq \exp\left[-\frac{\lambda a}{2}K(-a) - C(1+\lambda)\log(a)\right] \quad (23)$$

with an absolute constant  $C$ .

Now assume that  $\lambda_0 > 0$  is fixed, choose a small  $\hat{\varepsilon} > 0$  and assume that  $0 < \lambda \leq \lambda_0$ ,  $\lambda T \leq \hat{\varepsilon}$  and  $T \leq \hat{\varepsilon} \log(\frac{1}{\lambda T})$ . Set  $a = -K^{(-1)}(\frac{\lambda T}{4})$ . If  $\hat{\varepsilon}$  is chosen small enough then using the bounds of Proposition 6 one can readily show that

$$\frac{1}{2} \frac{\log(\frac{1}{\lambda T})}{\lambda T} \leq \sqrt{a} \leq 2 \frac{\log(\frac{1}{\lambda T})}{\lambda T} \quad (24)$$

which yields both  $\lambda\sqrt{a} \geq 2$  and  $a > 2$  (using  $\lambda \leq \lambda_0$ ) for small enough  $\hat{\varepsilon}$ . Thus we can use (23) to get

$$\begin{aligned}\log P(\tilde{\tau}_\lambda \leq T \left(1 + \frac{10T}{\log(1/(\lambda T))}\right)) &\geq -\frac{\lambda a}{2} \cdot \frac{\lambda T}{4} - C(1 + \lambda) \log a \\ &\geq -\frac{2}{T} \mathcal{W}^2(-\frac{\lambda T}{16}) - C \log(\frac{1}{\lambda T}) \\ &\geq -\frac{2}{T} \mathcal{W}^2(-\lambda T) - C \log(\frac{1}{\lambda T})\end{aligned}\quad (25)$$

where we again used Proposition 6 and the constant  $C$  can change from line to line (but only depends on  $\lambda_0$  and  $\hat{\varepsilon}$ ). Now set  $t = T \left(1 + \frac{10T}{\log(1/(\lambda T))}\right)$ , we can see that  $T \leq t \leq T(1 + 10\hat{\varepsilon})$ . The bound (25) gives

$$\begin{aligned}\log P(\tilde{\tau}_\lambda \leq t) &\geq -\frac{2}{t} \mathcal{W}^2(-\lambda T) \left(1 + \frac{10T}{\log(1/(\lambda T))}\right) - C \log(\frac{1}{\lambda T}) \\ &= -\frac{2}{t} \mathcal{W}^2(-\lambda T) - \frac{20T}{t} \mathcal{W}^2(-\lambda T) \frac{1}{\log(1/(\lambda T))} - C \log(\frac{1}{\lambda T}).\end{aligned}$$

Using the second bound of (15) (which is allowed since  $\lambda T \leq \hat{\varepsilon}$ ) together with  $1 \leq \frac{t}{T} \leq 1 + 10\hat{\varepsilon}$  we get that

$$\frac{20T}{t} \mathcal{W}^2(-\lambda T) \frac{1}{\log(1/(\lambda T))} + C \log(\frac{1}{\lambda T}) \leq C' \log(\frac{1}{\lambda t}).$$

Using the first and then the second bound of (15) we get

$$-\frac{2}{t} \mathcal{W}^2(-\lambda T) \geq -\frac{2}{t} (\mathcal{W}^2(-\lambda t) + C \log |\frac{T}{t}| \cdot \mathcal{W}(-\lambda t)) \geq -\frac{2}{t} \mathcal{W}^2(-\lambda t) - \frac{C'}{t} \log(\frac{1}{\lambda t}).$$

Putting everything together gives

$$\log P(\tilde{\tau}_\lambda \leq t) \geq -\frac{2}{t} \mathcal{W}^2(-\lambda t) - c \left(1 + \frac{1}{t}\right) \log(\frac{1}{\lambda t})$$

with a constant  $c$ . The only thing left to show is that for small enough  $\varepsilon' > 0$  if  $\lambda t \leq \varepsilon'$ ,  $t \leq \varepsilon' \log(1/(\lambda t))$  then we can find  $T$  with  $t = T(1 + \frac{10T}{\log(1/(\lambda T))})$  which satisfies  $\lambda T \leq \hat{\varepsilon}$  and  $T \leq \hat{\varepsilon} \log(\frac{1}{\lambda T})$ . This is fairly straightforward, one can check that if  $\varepsilon'$  is small enough then we can find a  $T$  with  $(1 - 10\varepsilon')t \leq T \leq t$  satisfying all the listed conditions.

Finally, to prove (17) for  $\tau_\lambda$  we use the comparison (9). Note that if we have the bound  $\lambda \leq \lambda_0$  then  $\lambda'' = \frac{\beta}{4} e^{-\frac{\beta}{4}t} \lambda \leq \lambda''_0 = \frac{\beta}{4} \lambda_0$ . By choosing a small enough  $\varepsilon$  in the bounds  $\lambda t \leq \varepsilon$  and  $t \leq \varepsilon \log(\frac{1}{\lambda t})$  we can use the just proved lower bound for  $\tilde{\tau}_{\lambda''}$  to give

$$\log P(\tau_\lambda \leq t) \geq -\frac{2}{t} \mathcal{W}^2(-\lambda \frac{\beta}{4} e^{-\frac{\beta}{4}t} t) - c \left(1 + \frac{1}{t}\right) \left(\log(\frac{1}{\lambda t}) + \frac{\beta}{4} t + \log \frac{4}{\beta}\right).$$

The fact that the right hand side is bounded from below by  $-\frac{2}{t} \mathcal{W}^2(-\lambda t) - \frac{c'}{t} \log(\frac{1}{\lambda t})$  with a new constant  $c'$  follows along the line of the already seen arguments using Proposition 6 and our assumptions on  $\lambda, t$ .  $\square$

Now we return to the proof of Lemma 9.

*Proof of Lemma 9.* Our main tool is again a coupling argument. We introduce the diffusions

$$\begin{aligned} dY_- &= \frac{\lambda}{2}\sqrt{\cosh^2 Y_- + a} dt - \frac{1}{2}dt + dB \\ dY_+ &= \frac{\lambda}{2}\sqrt{\cosh^2 Y_+ + a} dt + \frac{1}{2}dt + dB \end{aligned}$$

driven by the same Brownian motion as  $Y$ . Comparing the drifts of  $Y_-$ ,  $Y_+$  to that of  $Y$  we see that  $Y_- \leq Y$  and  $Y \leq Y_+$  until the blowup of the larger diffusions.

We also introduce  $Z_- = Y_- + t/2 - B_t$  and  $Z_+ = Y_+ - t/2 - B_t$  which satisfy the differential equations

$$Z'_- = \frac{\lambda}{2}\sqrt{\cosh^2(Z_- - t/2 + B_t) + a} > 0, \quad \text{and} \quad Z'_+ = \frac{\lambda}{2}\sqrt{\cosh^2(Z_+ + t/2 + B_t) + a} > 0$$

until their blowup times. Note that  $Z_-$  blows up exactly when  $Y_-$  blows up, and  $Z_+$  blows up when  $Y_+$  does. Thus  $Z_+$  will blow up before  $Y$  does and  $Z_-$  will blow up after.

We introduce the constants

$$\varepsilon_1 = \frac{10K(-a)}{\lambda\sqrt{a}}, \quad \varepsilon_2 = \frac{K(-a)}{8\lambda}, \quad \theta = \frac{4K(-a) + \varepsilon_1}{\lambda},$$

and we record for further use that since  $\lambda\sqrt{a} \geq 2$  we have

$$\varepsilon_1 > \frac{4\varepsilon_2 + \theta}{\sqrt{a}}. \tag{26}$$

We will show that on the event  $H = \{\sup_{t \leq \theta} |B_t| \leq \varepsilon_2\}$  the process  $Z_-$  will blow up before  $\theta$  and the process  $Z_+$  will blow up after  $\lambda^{-1}(4K(-a) - 2\varepsilon_1)$ . The final bound will come from an estimate on the probability of the event  $H$ .

Note that  $Z'_- > \frac{2\sqrt{a}}{\lambda} > 0$  until its blowup, in particular  $Z_-$  is strictly increasing. Moreover, on the event  $H$  for  $t \leq \theta$  we have

$$Z_-(t) - \theta/2 - \varepsilon_2 \leq Z_-(t) - t/2 + B_t \leq Z_-(t) + \varepsilon_2.$$

Let  $\eta_1$  be the hitting time of  $-\varepsilon_2$  by  $Z_-$ . Then for  $t \leq \eta_1 \wedge \theta$  on the event  $H$  we have

$$Z'_-(t) = \frac{\lambda}{2}\sqrt{\cosh^2(Z_-(t) - t/2 + B_t) + a} \geq \frac{\lambda}{2}\sqrt{\cosh^2(Z_-(t) + \varepsilon_2) + a}.$$

Rearranging, integrating and using identity (11):

$$\frac{\lambda}{2}(\eta_1 \wedge \theta) \leq \int_0^{\eta_1 \wedge \theta} \frac{Z'_-(t)}{\sqrt{\cosh^2(Z_-(t) + \varepsilon_2) + a}} dt \leq \int_{-\infty}^{-\varepsilon_2} \frac{1}{\sqrt{\cosh^2(z + \varepsilon_2) + a}} dz = K(-a),$$

which gives

$$\lambda(\eta_1 \wedge \theta) \leq 2K(-a) \leq \lambda\theta, \quad \text{and} \quad \lambda\eta_1 \leq 2K(-a).$$

Next we bound the travel time of  $Z_-$  from  $-\varepsilon_2$  to  $\theta/2 + \varepsilon_2$  which we denote by  $\eta_2$ . Since  $Z'_- \geq \frac{\lambda}{2}\sqrt{a}$  we have the bound

$$\lambda\eta_2 \leq \frac{4\varepsilon_2 + \theta}{\sqrt{a}} \leq \varepsilon_1.$$

Note that we have  $\eta_1 + \eta_2 \leq \lambda^{-1}(2K(-a) + \varepsilon_1) < \theta$  on  $H$ . Consider now the travel time  $\eta_3$  the from  $\theta/2 + \varepsilon_2$  to  $\infty$ . Denote by  $\hat{\eta}_3$  the minimum of  $\eta_3$  and  $\theta - \eta_1 - \eta_2 > 0$ . On the event  $H$  if  $Z_- \geq \theta/2 + \varepsilon_2$ ,  $\eta_1 + \eta_2 \leq t \leq \theta$  then we have the lower bound

$$Z'_-(t) = \frac{\lambda}{2}\sqrt{\cosh^2(Z_-(t) - t/2 + B_t) + a} \geq \frac{\lambda}{2}\sqrt{\cosh^2(Z_-(t) - \theta/2 - \varepsilon_2) + a}.$$

Rearranging the inequality and integrating from  $\eta_1 + \eta_2$  to  $\eta_1 + \eta_2 + \hat{\eta}_3$  we get

$$\lambda\hat{\eta}_3 \leq 2K(-a).$$

Collecting our estimates we see that on the event  $H$  we have

$$\lambda(\eta_1 + \eta_2 + \hat{\eta}_3) \leq 4K(-a) + \varepsilon_1 = \lambda\theta.$$

This immediately implies  $\hat{\eta}_3 = \eta_3$  and since  $\eta_1 + \eta_2 + \eta_3$  is the blowup time of  $Y_-$  we also get that  $\tau_{Y,\lambda,a} \leq \theta$  on  $H$ .

To show that the blow up time of  $Z_+$  is greater than  $\lambda^{-1}(4K(-a) - 2\varepsilon_1)$  on the event  $H$  we carry out a similar calculation. The process  $Z_+$  is strictly increasing and on the event  $H$  for  $t \leq \theta$  we have

$$Z_+(t) - \varepsilon_2 \leq Z_+(t) + t/2 + B_t \leq Z_-(t) + \theta/2 + \varepsilon_2.$$

Let  $\zeta_1$  be the hitting time of  $-\theta/2 - \varepsilon_2$  for process  $Z_+$ . Since the blowup time of  $Z_+$  is smaller than that of  $Z_-$  we get that  $\zeta_1 < \theta$  on the event  $H$ . For  $Z_+ < -\theta/2 - \varepsilon_2$  on  $H$  we get

$$Z'_+(t) = \frac{\lambda}{2}\sqrt{\cosh^2(Z_+ + t/2 + B_t) + a} \leq \frac{\lambda}{2}\sqrt{\cosh^2(Z_+(t) - \varepsilon_2) + a},$$

which gives us

$$\begin{aligned} \frac{\lambda}{2}\zeta_1 &\geq \int_0^{\zeta_1} \frac{Z'_+(t)}{\sqrt{\cosh^2(Z_+(t) - \varepsilon_2) + a}} dt = K(-a) - \int_{-\theta/2-2\varepsilon_2}^0 \frac{1}{\sqrt{\cosh^2(z) + a}} dz \\ &\geq K(-a) - \frac{\theta + 4\varepsilon_2}{2\sqrt{a}} \geq K(-a) - \varepsilon_1/2. \end{aligned}$$

In the last step we used (26). Now take  $\zeta_3$  to be the time spent by  $Z_+$  traveling from  $\varepsilon_2$  to  $\infty$ . Again note that the blow up time of  $Z_+$  is bounded by that of  $Z_-$  which is bounded by  $\theta$  on  $H$ . Therefore on  $H$  for  $\varepsilon_2 \leq Z_+$  we have

$$Z'_+(t) = \frac{\lambda}{2} \sqrt{\cosh^2(Z_+(t) + t/2 + B_t) + a} \leq \frac{\lambda}{2} \sqrt{\cosh^2(Z_+(t) + \theta/2 + \varepsilon_2) + a}.$$

This yields

$$\frac{\lambda}{2} \zeta_3 \geq K(-a) - \int_0^{\theta/2+2\varepsilon_2} \frac{1}{\sqrt{\cosh^2(z) + a}} dz \geq K(-a) - \varepsilon_1/2.$$

These estimates give

$$\lambda(\zeta_1 + \zeta_3) \geq 4K(-a) - 2\varepsilon_1.$$

Our last bound shows that on  $H$  the process  $Y$  blows up after  $\lambda^{-1}(4K(-a) - 2\varepsilon_1)$ . Putting everything together yields

$$P\left(\lambda\tau_{Y,\lambda,a} \in \left[4K(-a)\left(1 - \frac{5}{\lambda\sqrt{a}}\right), 4K(-a)\left(1 + \frac{5}{\lambda\sqrt{a}}\right)\right]\right) \geq P(H).$$

On the other hand we have

$$P(H) = P\left(\sup_{t \leq \lambda^{-1}\theta} |B_t| \leq \varepsilon_2\right) = P\left(\sup_{t \leq 1} |B_t| \leq \lambda^{1/2}\theta^{-1/2}\varepsilon_2\right)$$

and

$$\lambda^{1/2}\theta^{-1/2}\varepsilon_2 = \frac{K(-a)}{8\sqrt{4K(-a)\left(1 + \frac{10}{\lambda\sqrt{a}}\right)}} \geq \frac{1}{40}\sqrt{K(-a)}.$$

It is known that  $P(\sup_{t \leq 1} |B_t| \leq \delta) \geq 2(2\Phi(\delta/4) - 1)$  for  $\delta > 0$  (see e.g. Proposition 2.8.10 in [12]) which can be bounded from below by  $c\delta$  if we have an upper bound on  $\delta$ . Since  $a > 2$  we have an upper bound on  $\frac{1}{40}\sqrt{K(-a)}$  which finishes the proof.  $\square$

## 2.2 Proof of the upper bounds in Proposition 7

The proof of the upper bound will rely on the following estimate on the moment generating function of  $\tilde{\tau}_\lambda$ .

**Proposition 10** ([10]). *For any  $a > -1$  we have*

$$Ee^{-\frac{\lambda^2 a}{8}\tilde{\tau}_\lambda - \frac{\lambda\sqrt{|a|}}{4}\tilde{\tau}_\lambda} \leq e^{-\lambda((1+a)K(-a) - E(-a))}. \quad (27)$$

This is Proposition 8 in [10], note that in that statement  $-a$  is used in place of  $a$ .

*Proof of the upper bound in Proposition 7.* By Markov's inequality we have

$$P(\tilde{\tau}_\lambda \leq t) \leq (Ee^{A\tilde{\tau}_\lambda})e^{-At} \quad \text{for } A < 0.$$

We will choose  $A = -\frac{\lambda^2 a}{8} - \frac{\lambda\sqrt{a}}{4}$  with  $a = -K^{(-1)}(t\lambda/4) \geq -K^{(-1)}(\varepsilon/4) > 0$ . Using Proposition 10 we get

$$Ee^{A\tilde{\tau}_\lambda} \leq e^{-\lambda((1+a)K(-a)-E(-a))}.$$

This gives the upper bound

$$P(\tilde{\tau}_\lambda \leq t) \leq e^{-\lambda((1+a)K(-a)-E(-a))+\frac{\lambda^2 a}{8}t+\frac{\lambda\sqrt{a}}{4}t} \leq e^{K^{(-1)}(t\lambda/4)\frac{\lambda^2 t}{8}+\lambda E(-a)+\frac{\lambda\sqrt{at}}{4}},$$

where we used our specific choice of  $a$ . If  $\varepsilon$  is small enough we will have  $a \geq -K^{(-1)}(\varepsilon/4) > 2$  which allows us to use the asymptotics of Lemma 6. This gives the bound  $a \leq \frac{C}{\lambda^2 t^2} \log^2(\frac{1}{\lambda t})$  with a positive constant  $C$ , and also

$$\frac{K^{(-1)}(t\lambda/4)\lambda^2 t}{8} + \lambda E(-a) + \frac{\lambda\sqrt{at}}{4} \leq -\frac{2}{t}\mathcal{W}(-\lambda t/4)^2 + \frac{C}{t} + C(1 + \frac{1}{t})\log(\frac{1}{\lambda t})$$

(with a possibly different  $C$ ). Using the bounds in (15) and changing the value of  $C$  again we get

$$\frac{K^{(-1)}(t\lambda/4)\lambda^2 t}{8} + \lambda E(-a) + \frac{\lambda\sqrt{at}}{4} \leq -\frac{2}{t}\mathcal{W}(-\lambda t)^2 + C(1 + \frac{1}{t})\log(\frac{1}{\lambda t}),$$

which finishes the proof for  $\tilde{\tau}_\lambda$ .

To get the similar statement for  $\tau_\lambda$  we use (8) to obtain

$$P(\tau_\lambda \leq t) \leq P(\tilde{\tau}_{\frac{\beta}{4}\lambda} \leq t) \leq \frac{2}{t}\mathcal{W}^2\left(-\frac{\lambda\beta}{4}t\right) + c(1 + \frac{1}{t})\log\left(\frac{1}{\frac{\lambda\beta}{4}t}\right). \quad (28)$$

Note that we are allowed to use the just proved bound on  $\tilde{\tau}$  if we choose the upper bound  $\varepsilon$  on  $\lambda t$  small enough. To complete the proof we use (15) to replace  $\frac{\lambda\beta}{4}$  with  $\lambda$  everywhere by choosing  $\varepsilon$  small enough and modifying the constant in front of the last term of (28).  $\square$

### 3 Proof of the lower bound in Theorem 1

Our proof will have two main steps. In the first step we will give a reasonably good (but not optimal) lower bound on  $P(N_\beta(\lambda) \geq n)$ . In the second step we will show how one can get better and better bounds as we increase  $n$ .

*Proof of the lower bound in Theorem 1.* We first give a direct estimate on  $P(N_\beta(\lambda) \geq n)$  for  $\lambda \leq \lambda_0$ . Although this lower bound will be far from the statement of the theorem, we will later show how to ‘bootstrap’ this into a near optimal bound for large  $n$ .

Recall Theorem 2 and Proposition 3. The connection between the counting function  $N_\beta(\lambda)$  and the diffusion  $\alpha_\lambda(t)$  shows that for any  $T > 0$  we have

$$P(N_\beta(\lambda) \geq n) \geq P(\alpha_\lambda(T) \geq 2n\pi).$$

The drift in (3) on  $[0, T]$  is at least  $\lambda' = \frac{\beta}{4}e^{-\frac{\beta}{4}T}\lambda$  so coupling  $\alpha_\lambda$  and  $\tilde{\alpha}_{\lambda'}$  (using the same driving noise) we get

$$P(\alpha_\lambda(T) \geq 2n\pi) \geq P(\tilde{\alpha}_{\lambda'}(T) \geq 2n\pi).$$

For the process  $\tilde{\alpha}_{\lambda'}$  denote the hitting time of  $2\pi k$  by  $\tilde{\tau}_{\lambda',k}$  (note that  $\tilde{\tau}_{\lambda',1} = \tilde{\tau}_{\lambda'}$  using our previous notation). Using the same ideas needed for Proposition 3 it is easy to show that  $\tilde{\tau}_{\lambda',1}, \tilde{\tau}_{\lambda',2} - \tilde{\tau}_{\lambda',1}, \dots, \tilde{\tau}_{\lambda',k} - \tilde{\tau}_{\lambda',k-1}, \dots$  are i.i.d. random variables. Using this we can obtain the estimate

$$\begin{aligned} P(\tilde{\alpha}_{\lambda'}(T) \geq 2n\pi) &\geq P(\tilde{\tau}_{\lambda',1} \leq \frac{T}{n}, \text{ and } \tilde{\tau}_{\lambda',k} - \tilde{\tau}_{\lambda',k-1} \leq \frac{T}{n} \text{ for } 2 \leq k \leq n) \\ &= P(\tilde{\tau}_{\lambda'} \leq \frac{T}{n})^n \end{aligned}$$

We will set  $T = \frac{4}{\beta} \log(\frac{n}{\lambda})$  and apply the lower bound (17) of Proposition 7 with  $t = \frac{T}{n}$  and  $\lambda'$ . (Note that  $\lambda' \leq \frac{\beta}{4}\lambda_0$ .) We first have to check that the conditions of the proposition are satisfied. From the definitions we get

$$\lambda't = \left(\frac{\lambda}{n}\right)^2 \log\left(\frac{n}{\lambda}\right), \quad \frac{t}{\log\left(\frac{1}{\lambda't}\right)} = \frac{1}{n} \cdot \frac{4 \log\left(\frac{n}{\lambda}\right)}{2\beta \log\left(\frac{n}{\lambda}\right) - \beta \log \log\left(\frac{n}{\lambda}\right)}.$$

Since we have  $\lambda' \leq \frac{\beta}{4}\lambda_0$ , there is a positive integer  $n_0$  (depending only on  $\lambda_0, \beta$ ) so that if  $n \geq n_0$  then both of these expressions are bounded by  $\varepsilon$ . (Here  $\varepsilon$  is the constant from the statement of Proposition 7 which also only depends on  $\lambda_0, \beta$ .) Thus if  $n$  is large enough we can use Proposition 7 to get

$$\begin{aligned} P(\tilde{\tau}_{\lambda'} \leq \frac{T}{n}) &\geq \exp\left[-\frac{2}{t} \mathcal{W}^2(-\lambda't) - c\left(1 + \frac{1}{t}\right) \log\left(\frac{1}{\lambda't}\right)\right] \\ &= \exp\left[-\frac{\beta n}{2 \log\left(\frac{n}{\lambda}\right)} \mathcal{W}^2\left(\frac{\lambda^2}{n^2} \log\left(\frac{n}{\lambda}\right)\right) - c\left(1 + \frac{\beta n}{4 \log\left(\frac{n}{\lambda}\right)} \log\left(\frac{n^2}{\lambda^2} \cdot \frac{1}{\log\left(\frac{n}{\lambda}\right)}\right)\right)\right]. \end{aligned}$$

Choosing  $n$  large enough we can use the bounds (15) of Proposition 6 to get

$$P(\tilde{\tau}_{\lambda'} \leq \frac{T}{n}) \geq \exp\left[-c_1 n \log\left(\frac{n}{\lambda}\right) - c_2 n\right]$$

with positive constants  $c_1, c_2$  (which only depend on  $\lambda_0, \beta$ ). Putting our estimates together we obtain the lower bound

$$P(N_\beta(\lambda) \geq n) \geq \exp \left[ -c_1 n^2 \log\left(\frac{n}{\lambda}\right) - c_2 n^2 \right], \quad (29)$$

which holds for  $n \geq n_0$  with  $n_0$  only depending on  $\lambda_0, \beta$ . Finally, since  $P(N_\beta(\lambda) \geq k)$  is decreasing in  $k$ , we can adjust the constants  $c_1, c_2$  so that (29) remains valid for all  $n \geq 1$ .

To finish the proof we will show how to improve this lower bound to get the lower bound in Theorem 1. From the recursion equation of Lemma 4 we get that for any  $T > 0$  and  $n > 1$

$$\begin{aligned} P(N_\beta(\lambda) \geq n) &\geq E[\mathbf{1}(\tau_\lambda \leq T)g(\tau_\lambda)] \\ &\geq P(\tau_\lambda \leq T)g(T) = P(\tau_\lambda \leq T)P(N_\beta(\lambda e^{-\frac{\beta}{4}T}) \geq n-1), \end{aligned} \quad (30)$$

where we used the fact that the  $g(t)$  function is non increasing. Set  $T = \frac{4 \log(\frac{n}{\lambda})}{\beta n}$ . Using the same arguments as in the first part of this proof we can check that there is a positive integer  $n_0$  (depending only on  $\lambda_0, \beta$ ) so that for  $n \geq n_0$  we can use the lower bound of Proposition 7 with  $t = T$ . This gives

$$\log P(\tau_\lambda \leq T) \geq -\frac{\beta n}{2 \log(\frac{n}{\lambda})} \mathcal{W}^2 \left( -\frac{4}{\beta} \cdot \frac{\lambda}{n} \log\left(\frac{n}{\lambda}\right) \right) - c \left( 1 + \frac{\beta n}{4 \log(\frac{n}{\lambda})} \right) \log \left( \frac{\beta n}{4 \lambda} \cdot \frac{1}{\log(\frac{n}{\lambda})} \right).$$

Assuming that  $n$  is large enough we can use the bounds in (15) to remove the constant  $\frac{4}{\beta}$  from inside  $\mathcal{W}$ . This allows us to find positive constants  $c_1$  and  $c_2$  so that

$$\begin{aligned} \log P(\tau_\lambda \leq T) &\geq -\frac{\beta n}{2 \log(\frac{n}{\lambda})} \mathcal{W}^2 \left( -\frac{\lambda}{n} \log\left(\frac{n}{\lambda}\right) \right) - c_1 \log\left(\frac{n}{\lambda}\right) - c_2 n \\ &= -\frac{\beta}{2} n \log\left(\frac{n}{\lambda}\right) - c_1 \log\left(\frac{n}{\lambda}\right) - c_2 n \end{aligned} \quad (31)$$

where we used  $\mathcal{W}(x \log x) = \log x$  which holds for  $0 < x < 1/e$ .

Now assume that for a certain  $n$  there is a constant  $f_{n-1} > 0$  so that we have the lower bound

$$P(N_\beta(\lambda) \geq n-1) \geq \exp \left[ -f_{n-1} \log\left(\frac{n-1}{\lambda}\right) - c(n-1)^2 \right]$$

for all  $\lambda \leq \lambda_0$  with a fixed constant  $c$ . Then

$$\begin{aligned} \log P(N_\beta(\lambda e^{-\frac{\beta}{4}T}) \geq n-1) &\geq -f_{n-1} \log\left(\frac{n-1}{\lambda e^{-\frac{\beta}{4}T}}\right) - c(n-1)^2 \\ &= -f_{n-1} \log\left(\frac{n}{\lambda}\right) - f_{n-1} \log\left(\frac{n-1}{n}\right) - \frac{\beta}{4} T f_{n-1} - c(n-1)^2 \\ &\geq -\frac{n+1}{n} f_{n-1} \log\left(\frac{n-1}{n}\right) - c(n-1)^2. \end{aligned}$$

Together with (30) and (31) this yields the lower bound

$$\log P(N_\beta(\lambda) \geq n) \geq - \left( \frac{n+1}{n} f_{n-1} + \frac{\beta}{2} n + c_1 \right) \log\left(\frac{n}{\lambda}\right) - c_2 n - c(n-1)^2.$$

If we also assume that  $c > c_2$  then we can further simplify this to

$$\log P(N_\beta(\lambda) \geq n) \geq - \left( \frac{n+1}{n} f_{n-1} + \frac{\beta}{2} n + c_1 \right) \log\left(\frac{n}{\lambda}\right) - cn^2.$$

Thus we proved the following statement. Suppose that  $n_0$  is large enough,  $c > c_2$  and  $f$  solves the following recursion:

$$f_n = \frac{n+1}{n} f_{n-1} + \frac{\beta}{2} n + c_1, \quad \text{for } n > n_0. \quad (32)$$

Then if

$$\log P(N_\beta(\lambda) \geq n) \geq -f_n \log\left(\frac{n}{\lambda}\right) - cn^2$$

holds for  $n = n_0$  then it also holds for any  $n > n_0$ . Because of (29), we can choose  $f_{n_0}$  so that the previous inequality holds for  $n_0$  (with a suitable  $c$ ). To finish the proof, we need to show that the recursion (32) will have the appropriate asymptotics. The general solution of the recursion (without the initial condition) is

$$f_n = \frac{\beta}{2} n(n+1) + \left(c_1 - \frac{\beta}{2}\right) \sum_{k=1}^{n+1} \frac{n+1}{k} + C(n+1).$$

From this it is clear that there is a solution with any initial condition, and the solution satisfies

$$f_n \leq \frac{\beta}{2} n^2 + b n \log(n+1)$$

with a finite  $b$ . From this the lower bound of Theorem 1 follows for  $n \geq n_0$  and  $\lambda \leq \lambda_0$ . Using the argument following (29) we can now modify the constant  $b$  so that the lower bound holds for all  $n$  which completes the proof.  $\square$

## 4 Proof of the upper bound in Theorem 1

The upper bound will use a two step strategy similar to that of the lower bound, but our recursive step will be a bit more involved.

*Proof of the upper bound in Theorem 1.* Recall from Proposition 5 that

$$P(N_\beta(\lambda) \geq n) \leq \left( \frac{\lambda}{2\pi} \right)^n \leq \exp(-n \log(1/\lambda)).$$

By choosing  $c_1 > 1$  we get  $-n \log(1/\lambda) \leq -n \log(n/\lambda) + c_1 n^2$  which yields

$$\log P(N_\beta(\lambda) \geq n) \leq -n \log(n/\lambda) + c_1 n^2. \quad (33)$$

Now assume that  $n \geq 2$  and there are positive constants  $f_{n-1} \leq \frac{\beta}{2}(n-1)^2$ ,  $c_1$  so that for all  $0 \leq \lambda \leq \lambda_0$  we have

$$P(N_\beta(\lambda) \geq n-1) \leq \exp \left[ -f_{n-1} \log \left( \frac{n-1}{\lambda} \right) + c_1(n-1)^2 \right]. \quad (34)$$

Our goal is to show that we can choose an  $n_0$  (depending only on  $\lambda_0, \beta$ ) so that if  $n \geq n_0$  and  $c_1$  is larger than a fixed constant then we can get an upper bound of a similar form for  $P(N_\beta(\lambda) \geq n)$  as well.

Let  $T_0 = 0 < T_1 < \dots < T_M$ . Then

$$P(N_\beta(\lambda) \geq n) = \sum_{i=0}^{M-1} P(N_\beta(\lambda) \geq n, T_i \leq \tau_\lambda \leq T_{i+1}) + P(N_\beta(\lambda) \geq n, \tau_\lambda \geq T_M). \quad (35)$$

Choose real numbers  $\gamma_1 < \dots < \gamma_M$  with  $\gamma_1 = 1/3$ ,  $\gamma_M = 3$ ,  $M \leq 3n$  and  $\gamma_{i+1} - \gamma_i \leq 1/n$ . We set  $T_i = \gamma_i \frac{4 \log(n/\lambda)}{\beta n}$  for  $i = 1, \dots, M$ , if  $n \geq \lambda_0$  then these are positive numbers. We will estimate the sum (35) term by term.

Using the identity (5) from the proof of Lemma 4 we can show that

$$P(N_\beta(\lambda) \geq n, T_i \leq \tau \leq T_{i+1}) \leq P(N_\beta(\lambda e^{-\frac{\beta}{4}T_i}) \geq n-1) P(\tau < T_{i+1}). \quad (36)$$

From the definition of  $T_i$  (if  $n > \max(1, \lambda_0)$ ) we get

$$(1 + \gamma_i \frac{1}{n}) \log(n/\lambda) - \frac{1}{n-1} \leq \log \left( \frac{n-1}{\lambda e^{-\frac{\beta}{4}T_i}} \right). \quad (37)$$

Note that  $\lambda T_i \leq \frac{12}{\beta} \frac{\lambda}{n} \log(n/\lambda)$  which can be made as small we wish by choosing  $n_0$  large enough. Thus we can apply the upper bound in Proposition 7 to get

$$\begin{aligned} \log P(\tau < T_{i+1}) &\leq -\frac{2}{T_{i+1}} \mathcal{W}^2(-\lambda T_{i+1}) + c \left( 1 + \frac{1}{T_{i+1}} \right) \log \left( \frac{1}{\lambda T_{i+1}} \right) \\ &\leq -\frac{\beta}{2\gamma_{i+1} \log(n/\lambda)} \frac{n}{\lambda} \mathcal{W}^2 \left( \frac{\lambda}{n} \log \left( \frac{\lambda}{n} \right) \right) + C \left( 1 + \frac{\beta}{4\gamma_{i+1} \log(n/\lambda)} \frac{n}{\lambda} \right) \log \left( \frac{n}{\lambda} \right) \\ &\leq -\frac{\beta}{2\gamma_{i+1}} n \log \left( \frac{n}{\lambda} \right) + C \log \left( \frac{n}{\lambda} \right) + Cn. \end{aligned} \quad (38)$$

Here we used the bounds (15),  $\gamma_{i+1} \geq 1/3$  and the identity  $\mathcal{W}(x \log x) = \log x$ . The constant  $C$  depends only on  $\beta$  and  $\lambda_0$ , but may change from line to line.

Using the assumption (34) for  $\lambda e^{-\frac{\beta}{4}T_i} \leq \lambda_0$  we get

$$P(N_\beta(\lambda e^{-\frac{\beta}{4}T_i}) \geq n-1) \leq \exp \left[ -f_{n-1} \log \left( \frac{n-1}{\lambda e^{-\frac{\beta}{4}T_i}} \right) + c_1(n-1)^2 \right].$$

Using (37) we can write

$$\begin{aligned} -f_{n-1} \log \left( \frac{n-1}{\lambda e^{-\frac{\beta}{4}T_i}} \right) &\leq -f_{n-1} \left( 1 + \frac{\gamma_i}{n} \right) \log(n/\lambda) + \frac{f_{n-1}}{n-1} \\ &\leq - \left( f_{n-1} \left( 1 + \frac{\gamma_{i+1}}{n} \right) - \frac{\beta}{2} \right) \log(n/\lambda) + \frac{\beta}{2}(n-1), \end{aligned}$$

where we used  $f_{n-1} \leq \frac{\beta}{2}(n-1)^2$  and  $\gamma_{i+1} - \gamma_i \leq \frac{1}{n}$ , which holds for  $i \geq 1$ . We also have

$$\log \left( \frac{1}{\lambda T_{i+1}} \right) = \log(n/\lambda) - \log \log(n/\lambda) - \log(4\gamma_{i+1}/\beta) \leq \log(n/\lambda),$$

where we used  $1/3 \leq \gamma_i \leq 3$  and that  $n/\lambda$  can be made sufficiently large by choosing  $n_0$  appropriately.

Collecting all our estimates we get that for  $1 \leq i \leq M-1$  we have

$$\begin{aligned} \log P(N_\beta(\lambda) \geq n, T_i \leq \tau \leq T_{i+1}) &\\ &\leq - \left( f_{n-1} \left( 1 + \frac{\gamma_{i+1}}{n} \right) + \frac{\beta}{2\gamma_{i+1}} n \right) \log(n/\lambda) + c_1(n-1)^2 + C \log(\frac{n}{\lambda}) + Cn. \end{aligned}$$

Assuming  $c_1 > C$  this yields

$$\begin{aligned} \log P(N_\beta(\lambda) \geq n, T_i \leq \tau \leq T_{i+1}) &\\ &\leq - \left( f_{n-1} \left( 1 + \frac{\gamma_{i+1}}{n} \right) + \frac{\beta}{2\gamma_{i+1}} n - C \right) \log(n/\lambda) + c_1 n^2 - c_1 n \\ &\leq - \left( f_{n-1} + \sqrt{2\beta f_{n-1}} - C \right) \log(n/\lambda) + c_1 n^2 - c_1 n. \end{aligned} \tag{39}$$

In the last step we used that  $x(1 + \frac{\gamma}{n}) + \frac{\beta}{2\gamma} n \geq x + \sqrt{2\beta x}$  if  $x > 0, \gamma > 0$ .

For the  $i=0$  case we can use a similar argument (and the fact that  $\gamma_1 = \frac{1}{3}$ ) to show that

$$\log P(N_\beta(\lambda) \geq n, T_0 \leq \tau \leq T_1) \leq - \left( f_{n-1} + \frac{\beta}{2} 3n - C \right) \log(n/\lambda) + c_1 n^2 - c_1 n.$$

Since  $f_{n-1} \leq \frac{\beta}{2}(n-1)^2$  we have  $\frac{\beta}{2} 3n \geq \sqrt{2\beta f_{n-1}}$  and so inequality (39) holds for  $i=0$  as well.

To bound the last term in (35) we start with the bound

$$P(N_\beta(\lambda) \geq n, T_M \leq \tau) \leq P(N_\beta(\lambda e^{-\frac{\beta}{4}T_M}) \geq n-1).$$

Recalling that  $\gamma_M = 3$  and using arguments similar to what we used to obtain (39) we get

$$P(N_\beta(\lambda) \geq n, T_M \leq \tau) \leq -f_{n-1} \left(1 + \frac{3}{n}\right) \log(n/\lambda) + c_1 n^2 - c_1 n. \quad (40)$$

Applying the upper bounds (39) and (40) on the respective terms of the sum in (35) yields

$$\begin{aligned} \log P(N_\beta(\lambda) \geq n) &\leq -\min \left( f_{n-1} + \sqrt{2\beta f_{n-1}} - c_2, f_{n-1} \left(1 + \frac{3}{n}\right) \right) \log(n/\lambda) \\ &\quad + c_1 n^2 - c_1 n + \log(3n+1), \end{aligned}$$

with constants  $c_1, c_2$  that only depend on  $\lambda_0, \beta$ . By choosing  $c_1 > 2$  we can simplify this as

$$\log P(N_\beta(\lambda) \geq n) \leq -\min \left( f_{n-1} + \sqrt{2\beta f_{n-1}} - c_2, f_{n-1} \left(1 + \frac{3}{n}\right) \right) \log(n/\lambda) + c_1 n^2.$$

Combining this result with (33) we get the following statement. There is a positive integer  $n'_0$ , and positive constants  $c, c_1$  (all depending only on  $\lambda_0, \beta$ ) so that if we set  $f_{n_0} = n_0$  for an  $n_0 > n'_0$  and

$$f_n = \min \left( f_{n-1} + \sqrt{2\beta f_{n-1}} - c_2, f_{n-1} \left(1 + \frac{3}{n}\right), \frac{\beta}{2} n^2 \right), \quad \text{for } n > n_0, \quad (41)$$

then

$$\log P(N_\beta(\lambda) \geq n) \leq -f_n \log \left(\frac{n}{\lambda}\right) + c_1 n^2$$

for all  $n \geq n_0, \lambda \leq \lambda_0$ .

To finish our proof we just have to show that we can choose a large enough  $n_0$  as our starting point so that the solution of the recursion (41) satisfies  $f_n \geq \frac{\beta}{2} n^2 - cn \log(n+1)$  with a constant  $c$ .

This is fairly straightforward although somewhat technical. A simple computation shows that if we choose  $n_0$  large enough then the following statements hold:

- (i)  $f_n$  is strictly increasing,
- (ii)  $f_n = f_{n-1} \left(1 + \frac{3}{n}\right)$  for  $n_0 < n < n_1$  for a certain  $n_1$ ,
- (iii)  $f_n = \min \left( f_{n-1} + \sqrt{2\beta f_{n-1}} - c_2, \frac{\beta}{2} n^2 \right)$  for  $n \geq n_1$ , and
- (iv) if  $n$  is large enough and  $f_{n-1} \geq \frac{\beta}{2}(n-1)^2 - c(n-1) \log n$  with a large enough  $c$  then we also have

$$f_{n-1} + \sqrt{2\beta f_{n-1}} - c \geq \frac{\beta}{2} n^2 - cn \log(n+1).$$

Putting these statements together we get that  $f_n \geq \frac{\beta}{2}n^2 - cn \log(n+1)$  for a large enough  $n$  which gives

$$\log P(N_\beta(\lambda) \geq n) \leq -\left(\frac{\beta}{2}n^2 - cn \log(n+1)\right) \log\left(\frac{n}{\lambda}\right) + c_1 n^2 \quad (42)$$

for a large enough  $n$ . Using (33) we can now modify  $c, c_1$  so that (42) holds for all  $n$  and  $\lambda \leq \lambda_0$ , which finishes the proof.  $\square$

**Acknowledgements.** B. Valkó was partially supported by the National Science Foundation CAREER award DMS-1053280. The authors thank B. Virág for suggesting them the study of the overcrowding problem. They also thank P.J. Forrester, H. Nguyen and J. Yin for valuable comments.

## References

- [1] G. Anderson, A. Guionnet, and O. Zeitouni. *Introduction to random matrices*. Cambridge University Press, 2009.
- [2] G. Akemann, J. Ipsen and E. Strahov. Permanental processes from products of complex and quaternionic induced Ginibre ensembles. *Random Matrices: Theory Appl.* 3(4): 1450014, 2014.
- [3] G. Akemann and E. Strahov. Hole probabilities and overcrowding estimates for products of complex Gaussian matrices. *Journal of Statistical Physics* 151(6): 9871003, 2013.
- [4] P. Bourgade, L. Erdős, H.-T. Yau and J. Yin. Fixed energy universality for generalized Wigner matrices. [arXiv:1407.5606](https://arxiv.org/abs/1407.5606), 2014.
- [5] E. L. Basor, C. A. Tracy, and H. Widom. Asymptotics of level-spacing distributions for random matrices. *Phys. Rev. Lett.*, 69(1):5–8, 1992.
- [6] F. J. Dyson. Statistical theory of energy levels of complex systems II. *J. Math. Phys.*, 3: 157–165, 1962.
- [7] L. Erdős, B. Schlein and H.-T. Yau. Wegner Estimate and Level Repulsion for Wigner Random Matrices. *International Mathematics Research Notices*, 3: 436–479, 2010.
- [8] M. M. Fogler and B. I. Shklovskii. Probability of an eigenvalue number fluctuation in an interval of a random matrix spectrum. *Phys. Rev. Lett.*, 74(17):3312–3315, Apr 1995.

- [9] P. J. Forrester. *Log-gases and random matrices*, volume 34 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2010.
- [10] D. Holcomb and B. Valkó. Large deviations for the Sine $_{\beta}$  and Sch $_{\tau}$  processes. *Probab. Theory Related Fields*, 2014. DOI 10.1007/s00440-014-0594-4
- [11] M. Krishnapur. Overcrowding estimates for zeroes of Planar and Hyperbolic Gaussian analytic functions. *Journal of Statistical Physics*, (124), No. 6, 1399–1423, 2006.
- [12] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.
- [13] M. L. Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, 2004.
- [14] B. Valkó and B. Virág. Continuum limits of random matrices and the Brownian carousel. *Inventiones Math.*, 177:463–508, 2009.
- [15] B. Valkó and B. Virág. Large gaps between random eigenvalues. *Ann. Probab.*, 38(3):1263–1279, 2010.